Data-analytical Stability in Second-Level fMRI inference

Sanne P. Roels, Tom Loeys, & Beatrijs Moerkerke
Department of data analysis, Ghent University
sanne.roels@ugent.be

problem

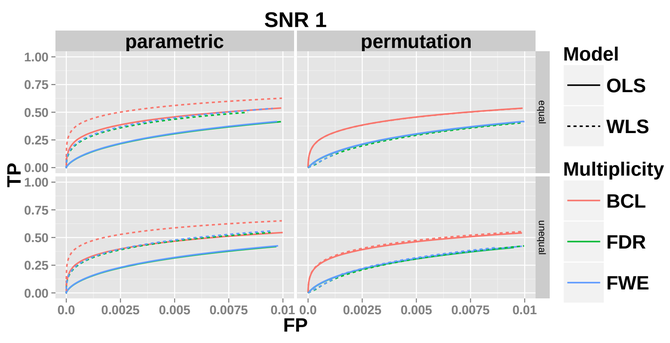
- ◇ **fMRI** plays an important role to localize brain regions
- ◇ 2-level **mass-univariate** GLM-based approach still popular for multi-subject analysis
- ◇ in voxel-based analysis: 1) continued use of **different models** to pool information over subjects; 2) **various multiple testing** procedures to control false positives and 3) **distinct inference procedures**
- ◇ choices in each of three phases impact reproducibility
- ◇ today no golden standard for a voxel-based approach
- ◇ **data-analytical stability** has the ability to further inform these choices, while it is also a proxy for **reproducibility** and intrinsically allows for an evaluation on real data

model - inference - multiple testing

- ◇ from **1st level model** the contrast image $\boldsymbol{c\hat{\beta}}$ from $v$oxel-wise model $\boldsymbol{Y_v} = \boldsymbol{X\beta_v} + \boldsymbol{\epsilon_v}$ is fed into
- ◇ **the 2nd level GLM**[1]: yet no consensus on how to pool information over subjects [2, 3]
  - **OLS** ignores subject-specific variability (only uses $\boldsymbol{c\hat{\beta}}$)
  - **WLS** accounts for subject specific variability (also includes the use of $var(\boldsymbol{c\hat{\beta}})$) via down-weighting
- ◇ **account for multiple testing**: also no standard but choice between theory-driven [4] and practical approaches [5] for **inference** based on parametrical assumptions or on permutations [3]
  - **FWE:** control probability of at least 1 False Positive (FP) (via RFT (parametrical inference) or max statistic (permutation))
  - **FDR:** control proportion of FPs over all selected voxels (via p-values that are computed via parametrical inference or permutation)
  - **BCL:** control FPs by setting a voxel-wise (**B**onferroni-like) threshold and specify a minimal **CL**uster size (here 10)

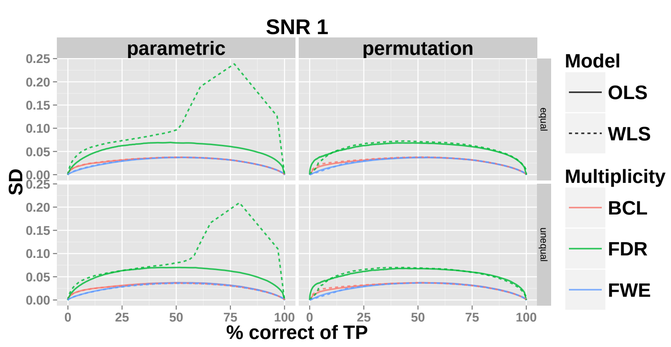
simulation study

balance FP and TP



- ◇ WLS outperforms OLS
- ◇ BCL outperforms FDR and FWE
- ◇ permutation no diff. between OLS and WLS
- ◇ for higher SNR, smaller differences

stability



- ◇ FDR is always more variable
- ◇ BCL comparable with FWE
- ◇ permutation ≈ parametrical inference
- ◇ high SNR, FDR + permutation: very instable

goals

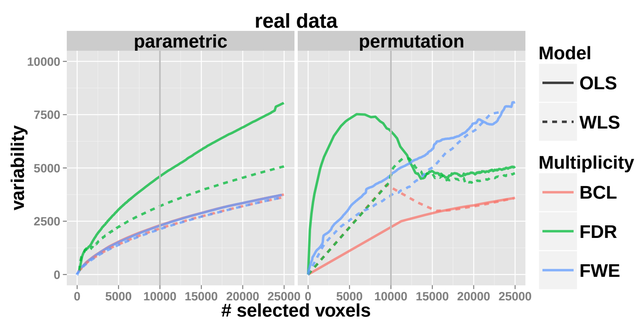
- ◇ **comparison** of combinations of methods for analysis of multi-subject fMRI
- ◇ to achieve a **golden standard** for analysis
- ◇ illustrate how to evaluate based on both **simulation and real data**

methods

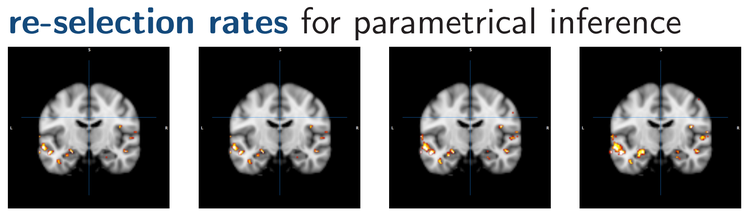
- ◇ **simulation** 2 clusters in $45 \times 45 \times 45$ volume; no first-level noise; smoothing applied to comply RFT assumptions; 2 SNR settings
- ◇ **real data**: 15 subjects from HCP data
- ◇ **criteria**
  - **type I&II error rates** Receiver-Operator-Curve
    False and True Positives (**FP**,**TP**) versus False and True Negatives (**FN**,**TN**)
  - **stability** variability on the number of selected voxels and *re-selection rate*: at a fixed threshold compute the # a voxel is selected over K bootstrap samples
- ◇ **procedures** `FSL`-based, in-house `R`-based

real data

◇ stability



- ◇ **re-selection rates** for parametrical inference



FDR-OLS FDR-WLS FWE-OLS FWE-WLS

- ◇ FDR: more variable + has lower re-selection rates
- ◇ WLS is found to be less variable

discussion and conclusions

- ◇ **model** WLS is slightly more stable than OLS, in ROC WLS outperforms OLS
- ◇ **multiple testing correction** FWE and BCL are almost always more stable than FDR, in ROC BCL outperforms FDR/FWE
- ◇ **inference** both for FP/TP balance and stability we find no large differences between permutation-based and parametrical inference

→ with respect to FP/TP balance, BCL offers a good alternative, but lacks statistical justification

→ with respect to stability, FDR is outperformed by the other procedures

references

[1] Beckmann,et al. (2003) General multi- level linear modeling for group analysis in fMRI,*NeuroImage*
[2] Mumford and Nichols (2009). Simple group fMRI modeling and inference, *NeuroImage*
[3] Thirion, et al. (2007). Analysis of a large fMRI cohort:. . . . *NeuroImage*
[4] Bennett et al. (2009). The principled control of false positives in neuroimaging. *SCAN*
[5] Lieberman et al. (2009) Type I and Type II error concerns in fMRI research: . . . *SCAN*

Acknowledgements: This work was carried out using the Stevin Supercomputer Infrastructure at Ghent University, funded by Ghent University, the Hercules Foundation and the Flemish Government - department EWI.

inter@ctive

`http://tinyurl.com/j9qsjjd`